Benchmarking computational methods for crystal structure prediction using experimental enantiomer excess and precipitation data

N. Vides[1], E. J. Blancas[1], J. A. Montero-García[1], J. M. Costa-Fernandez[1], V. del Amo[2], C. Concellón[2] and A. Otero-de-la-Roza[*1]

[1]Department of Physical and Analytical Chemistry, University of Oviedo, C/ Julián Clavería 8, 33006 Oviedo, Spain,
[2]Department of Organic and Inorganic Chemistry, University of Oviedo, C/ Julián Clavería 8, 33006 Oviedo, Spain

Email of communicating: videsnuria@uniovi.es

The accurate prediction of the experimentally observable crystal structure from the molecular diagram, a field known as *crystal structure prediction* (CSP), remains a central challenge in computational chemistry [1]. CSP is intimately related to the study of polymorphism in organic molecular crystals, and has industrial and technological applications, particularly in the design of pharmaceutical formulations, agrochemical products, and organic electronic materials [2]. Ranking functions are crucial components of CSP protocols, comprising one or more computational approaches used to calculate relative lattice energies between candidate crystals. Since the success of the CSP protocol is closely tied to the accuracy of ranking methods, benchmarking their performance is essential [3]. However, this is difficult due to the lack of reference benchmark data in periodic crystals. In this work, we present a novel approach to benchmarking CSP methods using experimental molecular crystal solubility data. We developed and implemented an experimental protocol to extract thermodynamic free energy differences (ΔG) between crystal forms from solubility data, enabling the generation of high-quality benchmark data suitable for assessing the performance of various computational ranking functions. The relevance of the experimental ΔG is evaluated based on theoretical considerations and control experiments. The experimental data generated in this work is used to assess DFT-based computational methods, including vibrational contributions to the free energy. In particular, we examine how *anharmonic effects* impact the calculation of relative free energies between crystal forms using our new implementation of Allen's quasiparticle theory [4]. Our results highlight the importance of appropriate benchmarking of CSP ranking methods and establish experimental solubility-derived equilibrium data as a robust and practical reference for validating computational methodologies. This work provides both a rigorous benchmarking framework and new insights into the thermodynamic modelling of molecular crystalline systems.

[1] Hoja, J., Ko, H.-Y., Neumann, M. A., Car, R., DiStasio, R. A., & Tkatchenko, A. (2019). *Sci. Adv*., **5(1)**, eaau3338.

[2] Beran, G. J. O. (2023). *Chem. Sci.*, **14**, 13290-13312.

[3] Hunnisett, L. M., Francia, N., Nyman, J., Abraham, N. S., Aitipamula, S., Alkhidir, T., Almehairbi, M., Anelli, A., Anstine, D. M., Anthony, J. E. et al. (2024). *Acta Crystallogr. B*, **80(6)**, 548-574.

[4] Blancas, E. J., Lobato, A., Izquierdo-Ruiz, F., Márquez, A. M., Recio, J. M., Nath, P., Plata, J. J., and Otero-de-la-Roza, A. (2024) *npj Comput. Mater.*, **10**, 267.